

# Worst-Case Optimal Battery Filling Policies with Constrained Adjustable Service

Jean-Yves Le Boudec, *Fellow, IEEE*, Dan-Cristian Tomozei, *Member, IEEE*

**Abstract**—We study battery filling policies with hard deadlines when the supply of energy can be modulated. This occurs for example with an electric plug-in vehicle using an adjustable electricity service for charging; such a service is offered in some countries as a means to provide flexibility to operators, and typically involves non-scheduled service reductions combined with a service guarantee that constrains these reductions. The problem for the battery user is to determine a charging policy, which we call a “consumption policy”, that meets a given “full battery” deadline while minimizing the energy cost (i.e. the bill paid to the electricity provider). As the charging efficiency is diminishing with respect to consumption, it is not optimal to charge as much and as early as possible. On the other hand service reductions cannot be predicted but it is possible to gain some information on the worst case reduction by analyzing past reductions. In this context, the computation of a causal consumption policy is an open problem. In this paper we consider a battery user interested in charging her battery while minimizing the worst case cost, where the total cost is a sum of two terms that reflect (i) the total energy consumption and (ii) the distance to a full battery at the deadline. We prove that there exists a causal consumption policy that minimizes the worst-case cost of the user. We find that the policy is threshold-based and give an efficient method to explicitly compute it at any time based solely on knowledge of past reductions, of the service guarantees and on the current distance to completion. Our method is based on the use of service curves and game theory.

**Index Terms**—Service Curves; Battery Management; Adjustable Supply; Repeated Game; Worst-Case Cost.

## I. INTRODUCTION

### A. Flexible Electricity Services

Flexible electricity services are one of the tools used by utilities to control energy flows and cope with variability in demand and supply. Flexible services can use control by price or by quantity. Control by price has been extensively studied and convergence to an optimal allocation has been showed when users maximize a utility function that depends on the shed load (e.g., [4], [8]). In this paper, we are interested in the latter case, more precisely in *adjustable services*. Such services are successfully performed by many companies (Romande Energie, Peakshaver, Voltalis): for example, Voltalis controls residential electrical heating via on-off signals and guarantees a heating interruption of at most one hour per day; similarly, Romande Energie offers, in addition to a non controlled service, an interruptible service where supply can be interrupted for at most 4 hours every day [1], [2], [9], recommended for water boilers, or electrical heating.

Adjustable services typically target large, flexible loads, that have a significant impact on the energy flows in the grid. Consumers adhere to such adjustable services because the rate per kWh is much lower than that of regular full service contracts. In addition to the traditional thermal loads, plug-in electric vehicles (PEVs) are excellent candidates for such adjustable services: they are large, flexible consumers. Moreover, an important part of the cost associated with the EV is refueling, and it is always interesting to have a lower rate for battery recharging. From the grid operator perspective, EVs strain the electrical grid, so it is also important to be able take advantage of their flexibility.<sup>1</sup>

A suitable way to model “adjustable contracts” is via a *service curve* contract [7]. Such a contract stipulates that the consumer is guaranteed an amount of energy of at least  $\beta(t)$  kWh over *any* time window of length  $t$  hours, where  $\beta$  is a non-decreasing function called “the service curve”. In particular, we consider a so-called “load-switching” service curve, defined in detail in Section II-A, which allows unrestricted consumption up to a maximum power for at least a predetermined fraction  $f$  of each day, and thus service interruptions during at most  $(1 - f)$  of each day.

### B. Flexible Battery Charging

We consider a consumer who owns a battery that needs to be charged before a given deadline. In addition to the standard electricity contract that guarantees full service availability, she has agreed with the utility on a service curve contract that provides an adjustable service, adapted for battery charging (large maximum power). This second contract specifies a flat rate per kWh, typically much lower than that of the standard contract, and a “load-switching” curve, similar to the one currently used for water boilers or electrical resistive heaters. The consumer has access to a *differentiated service*: the battery can be plugged either on the “adjustable service” plug (cheaper), or on the regular plug (much more expensive).

We assume that the consumer takes advantage of the differentiated service. For this, she has selected the “load-switching” contract that is least inconvenient to her in the long run: she should not be affected by service reductions and should have a full battery at the deadline (hence the flexibility of the charging process). She thus wants to have the illusion of an “always available” system, with transparent service reductions. Such contracts are very interesting for companies, such as the Post

<sup>1</sup>It is estimated that a transition from fossil-fuel powered vehicles to PEVs entails a significant increase in global electricity consumption (roughly from simple to double). Such a transition would require important investments in the electrical grid infrastructure, as it is currently unable to support such a dramatic increase.

Office, who operate a large fleet of PEVs during well defined time periods (such as, say, from 6am to 6pm). Moreover, individual PEV owners can also benefit from a well tailored adjustable service, despite being affected by “range anxiety”: they are always given the option of advancing the charging deadline, or, in exceptional situations, connecting the battery to the regular plug, thus getting full service (at a higher price).

In this paper we assume that the energy flow is from the grid to the battery, i.e., the battery is not feeding energy into the grid (the other cases are for further study). Thus the battery appears to the grid as an elastic load.

To summarize, on a day-to-day basis, the consumer uses the adjustable service defined by a service curve and wishes to have her battery fully charged at the deadline while paying as little as possible to the electricity provider. We model the consumer cost as the sum of (i) the amount of money paid for the service at a flat rate (e.g., the electricity bill) and (ii) the disutility of a discharged battery expressed as the distance to completion at the deadline (e.g., how far the battery is from being full at the deadline). We choose weights for these two terms such that the second term dominates the first, i.e., the minimum cost consumption always achieves having the battery closest to full at the deadline.

The consumer is risk-averse: by observing past reductions and through knowledge of the service curve, she wishes to protect against worst-case future reduction signals and thus to minimize the worst-case cost.<sup>2</sup> Here is where the service curve plays an important role: we show that the worst-case signals can be anticipated. In particular, if the worst-case reduction means missing the deadline, then thanks to the chosen cost function, the consumer will do everything in her power to have a battery as close to full as allowed by the service reductions.

A naive approach for such a consumer consists in always consuming as much as allowed, until the consumption needs are satisfied. However, the marginal advancement of state of charge is decreasing with respect to the instantaneous consumed power [6], [10], due to inefficiencies and thermal losses (which are quadratic with respect to current), so that the speed at which the battery gets charged increases sublinearly with the instantaneous power consumption. Therefore, the naive approach is wasteful and is generally not optimal. In contrast, an alert battery user can keep a record of past service reductions and infer bounds on future reductions, thanks to the service curve guarantee. In this context, the computation of an optimal causal consumption policy is an open problem.

### C. Contributions

We prove that there exists a causal consumption policy that minimizes the worst-case cost of the consumer. We find that the policy is threshold-based and can be explicitly computed: at a given time the optimal consumption is determined as the minimum of the allowed consumption and a certain adaptive threshold that depends solely on the knowledge of past reductions, on the service curve, and on the current completion level. Incidentally, we find that the optimal policy corresponds to a

subgame-perfect Nash equilibrium: our method of proof shows that, at each time step, the policy that minimizes the worst-case cost corresponds to the Nash equilibrium of a two-player zero-sum game played by the consumer and the distribution system operator, and we give a general method for constructing an optimal causal consumption policy based on refining such Nash equilibria to include decisions taken at each time step (like in a multi-stage game).

*Outline:* In Section II-A we describe the service curve contract that binds the distribution system operator and the consumer. In Section II we describe formally the consumer problem and we give the main result (the optimal consumption policy) in the form of Algorithm 1. Sections IV and V are dedicated to the proof of this result. In Section III-C we evaluate numerically the optimal consumption policy, the omniscient, and the greedy one. We conclude in Section VI.

## II. SYSTEM MODEL AND DEFINITION OF OPTIMAL POLICY

In this section we first we describe the system model and the adjustable service contract; then we give a model of the battery (which can easily be generalized to other energy-hungry tasks involving storage systems). Next, we explicitly describe the cost of the consumer as previously motivated. Last, we define worst-case optimal policies.

*Notation:* We introduce various notations which we summarize in Table I. We also use the following convention. For any  $s \leq t$  and a sequence  $\vec{x} \stackrel{\text{def}}{=} (x_t)_t$ , we denote by  $\vec{x}_{:t}$  the sequence  $(x_{t'})_{t' \leq t}$  and by  $\vec{x}_{s:t}$  the sequence  $(x_{t'})_{s \leq t' \leq t}$ . For some constant  $c$ , the notation “ $\vec{x} \leq c$ ” means that  $x_t \leq c$  for all  $t$ . Moreover, given two sequences  $\vec{x}$  and  $\vec{y}$ , we denote the concatenation of the two by  $[\vec{x} \vec{y}]$ .

### A. Adjustable Service

We consider a consumer that uses an adjustable service to charge a battery by a deadline  $T$ , by drawing electricity provided by a “Distribution System Operator” (DSO). We use the concept of service curve contract [7] to model service guarantees provided by the DSO for the adjustable service. For simplicity we consider that time is discrete and slotted. At every time slot, the DSO sends a control signal  $x_t$  to the consumer. Let  $z_t$  be the net energy drawn by the consumer from the grid in time slot  $t$  (in units of energy). The contract between the consumer and the distribution system operator specifies the following constraints. First,

$$0 \leq z_t \leq x_t \leq \bar{x} \text{ for all } t \quad (1)$$

i.e., the consumed energy may not exceed the level  $x_t$  specified by the DSO<sup>3</sup>;  $\bar{x}$  is the maximum that may ever be allowed.

Second, the consumer receives the guarantee that throttling cannot be excessive. This can be expressed by requiring

$$\sum_{\tau=t'+1}^t x_\tau \geq \beta(t-t') \text{ for all } t', t \text{ with } t' < t, \quad (2)$$

<sup>2</sup>This can be done automatically by a device that controls the charging process based on monitoring the state of the battery and past reduction signals.

<sup>3</sup>We assume  $z_t$  is positive since there is no battery-to-grid flow.

TABLE I  
NOTATION LIST

DSO	Distribution System Operator (electricity provider)
$x_t$	control signal sent by DSO, maximum amount of energy that can be drawn in time slot $t$
$z_t$	energy consumed by consumer in order to charge her battery in time slot $t$ ; cannot exceed $x_t$
$\pi$	used to denote a consumer policy, <i>i.e.</i> , a mapping between a sequence of control signals $\vec{x}$ and corresponding charging decisions (energy consumption) $\vec{z}$ , Section II-D
$B(\vec{z}_{1:t}) = B_0 + \sum_{s=1}^t f(z_s)$	battery level at the end of time slot $t$ given feasible consumptions $\vec{z}$
$T$	deadline at which the consumer wishes to have her battery fully charged
$p$	the price of energy bought by the consumer
$K$	factor of disutility of discharged battery at the deadline $T$
$f(z)$	increase in battery level when consuming $z$ units of energy, Eq.(6)
$\phi(z)$	$\stackrel{\text{def}}{=} pz - Kf(z)$ , the increase in the cost of the consumer when consuming $z$ units of energy, Section II-C
$c(\cdot, \cdot)$	$c(\pi, \vec{x}) \stackrel{\text{def}}{=} \sum_{t=1}^T \phi(z_t)$ , the cost of policy $\pi$ under control $\vec{x}$
$\beta(\cdot)$	service curve; in this paper we use load-switching curves (3) with parameters $t_0$ , $t_1$ , and $\bar{x}$ (see below)
$t_0, t_1$	the load-switching service curve guarantees maximum service interruption of $t_0$ during any time window of length $t_1$
$\bar{x}$	maximum amount of energy that can ever be consumed in one time slot

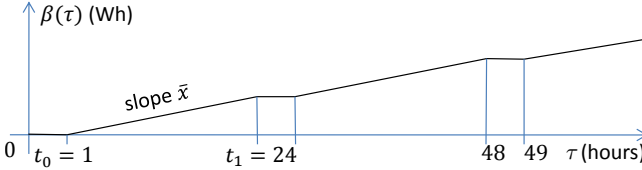


Fig. 1. The “load-switching” service curve  $\beta$ , shown here, allows the DSO to switch off the load for at most 60 minutes every day; it also allows more complex control signals, such as, for example, reducing the load to  $\bar{x}/2$  for 120 minutes every day.

where  $\beta : \mathbb{N} \rightarrow [0, +\infty)$  is a function defined at contract subscription time, called a *service curve*. Specifically, we take  $\beta$  to be a *load-switching* service curve with parameters  $t_0$ ,  $t_1$ , and  $\bar{x}$ , *i.e.*, the function illustrated in Figure 1 and defined by

$$\beta(t) = t_0 \bar{x} b_{t_1/t_0}(t/t_0), \quad (3)$$

where  $b_s(x) \stackrel{\text{def}}{=} \max\{x, \lfloor x/s \rfloor s + 1\} - (\lfloor x/s \rfloor s + 1)$ . This specific choice of service curve corresponds to a service that may be interrupted at most  $t_0$  time units every  $t_1$  time units; in particular, the electricity contracts currently used for boilers and heating [1], [2], [9] can be modeled this way.

It was shown [7, Theorem 1] that a sequence of control signals  $\vec{x} \stackrel{\text{def}}{=} (x_t)_t$  satisfies the service curve constraint in (2) if and only if  $\bar{x} \leq \bar{x}$  and the allowed energy consumption over any time window of length  $t_1$  is at least  $\beta(t_1)$ . In other words, the structure of the function  $\beta$  allows us to replace the condition (2) by the simpler one:

$$\sum_{\tau=t+1}^{t+t_1} x_\tau \geq \beta(t_1) \text{ for all } t. \quad (4)$$

We say that a sequence of control signals  $\vec{x}$  is *feasible* if it satisfies (1) and (2) or, equivalently, (1) and (4).

### B. Storage Model

The consumer wants to charge a battery before a given deadline  $T$  without knowledge of the future signals  $x_1, \dots, x_T$ . At the end of time slot  $t$  the battery state is characterized by a storage level  $0 \leq B(t) \leq \bar{B}$ , where  $\bar{B}$  is the energy capacity of the battery<sup>4</sup>. The consumer arrives at the end of time slot  $t = 0$  with a battery in state  $B(0) = B_0$ . She needs to immediately choose a charging strategy that ensures (if possible) a full battery at the end of time step  $T$ . To this end, she uses the knowledge of the service curve, of the past control signals, and of the battery characteristics.

The battery charging procedure consumes an amount of energy  $z_t$  during slot  $t$ , with  $0 \leq z_t \leq x_t \leq \bar{x}$ , as imposed by (1), where  $\bar{x}$  is the maximum allowed. By the end of time slot  $t$ , the battery level increases by

$$\Delta B(t) \stackrel{\text{def}}{=} B(t) - B(t-1) = f(z_t),$$

where  $f : [0, \bar{x}] \rightarrow [0, \infty)$  is a differentiable, concave, increasing function, that satisfies  $f(0) = 0$  (we do not consider battery leakage, a battery will not lose charge if it is left alone). The charging stops when the battery is full. Hence,  $z_t$  must always be chosen such that there is no overflow:  $B(t) \leq \bar{B}$ , or equivalently, since  $B(t) = B(t-1) + f(z_t)$ ,

$$f(z_t) \leq \bar{B} - B(t-1) \text{ for all } t. \quad (5)$$

We prolong the inverse  $f^{-1} : [0, \infty) \rightarrow [0, \bar{x}] \cup \{+\infty\}$ : for all  $y > f(\bar{x})$  we set  $f^{-1}(y) = +\infty$ .

Concretely, one can consider a charging efficiency  $\eta$  and an internal resistance  $\rho$  (which accounts for thermal losses). Then the increase in battery level is given by

$$f(z) = \eta z - \rho z^2. \quad (6)$$

Here  $\bar{x}$  must satisfy  $\bar{x} < \frac{\eta}{2\rho}$  (otherwise  $f$  is not increasing).

### C. Cost Function

The total cost incurred by the customer is a sum of her consumption cost at a flat rate  $p$  (*i.e.*, the electricity bill) and a linear term reflecting the disutility of having a discharged battery at the deadline  $T$ . It can be written as  $\sum_{t=1}^T p z_t + K(\bar{B} - B(T))$ , where  $K$  is a positive constant. We must pick  $K$  “large enough”, *i.e.*,  $K \geq \frac{p}{f'(\bar{x})}$ . This is needed to ensure that the battery level attained at the deadline by the consumption sequence of minimum cost is always the highest among all other consumption sequences. We give more precise arguments below.

Denote the cost increment at each time step by

$$\phi(z) \stackrel{\text{def}}{=} pz - Kf(z).$$

<sup>4</sup> $B(t)$  represents a scaling of the actual state of charge, since EV batteries are not discharged below a certain minimal level.

Consider only the controllable cost of the consumer (*i.e.*, the part of the cost that depends on  $\vec{z}$ ); it can be expressed as  $C(\vec{z}_{1:T})$ , or in short  $C(\vec{z})$ , where

$$C(\vec{z}_{t_1:t_2}) = \sum_{s=t_1}^{t_2} \phi(z_s). \quad (7)$$

Consider a hypothetical omniscient consumer who knows the future in advance. For a given control sequence  $\vec{x} = (x_t)_{t=1}^T$  she can determine an optimal consumption sequence:  $\omega^*(\vec{x}) = \vec{z}^*$  such that

$$\vec{z}^* \in \arg \min \{C(\vec{z}) : \vec{z} \leq \vec{x} \text{ and } \vec{z} \text{ satisfies (5)}\}. \quad (8)$$

Since we consider risk-averse consumers (that care more about having a full battery at the end of the day than about the electricity bill), we want the following property verified:

For any valid consumption sequence  $\vec{z}$ , *i.e.*, that satisfies (1) and (5), the level at the deadline of a battery charging according to  $\vec{z}$  is less or equal to that of the same battery charging according to  $\vec{z}^*$ .

When  $\phi$  is strictly decreasing, this is readily verified. Indeed, if with  $\vec{z}^*$  the battery is full at the deadline, then the property holds. Otherwise, if the level at the deadline with  $\vec{z}^*$  is  $B^* < \bar{B}$ , then it must be that  $\vec{z}^* = \vec{x}$ . Indeed, otherwise there exists a time slot  $t$  when  $z_t^* < x_t$ , and by increasing consumption at this time step by  $0 < \epsilon < \min\{x_t - z_t^*, f^{-1}(\bar{B} - B^*)\}$  we get a better filled battery and a lower cost, since  $\phi(z_t^*) > \phi(z_t^* + \epsilon)$ .

If  $K$  is too small, *i.e.*,  $K \leq \frac{p}{f'(0)}$ , then  $\phi$  is increasing, and the optimal choice is to always consume nothing. Intuitively, the electricity bill matters much more than having a full battery, so the consumer never consumes. When  $K$  is between the two values, *i.e.*,  $\frac{p}{f'(0)} \leq K \leq \frac{p}{f'(\bar{x})}$ , the consumer might choose to charge less in order to get a smaller bill, and the desired property again does not hold.

Note: In addition to a larger electricity bill, charging at high power is undesirable in general since it may cause undercharging: a battery that is charged at high power reaches too fast the upper voltage limit beyond which the battery electrodes may be damaged, and the charging process is stopped early. However, in this paper we consider EV batteries that use a relatively small depth-of-discharge, as is common for hybrid vehicles (*i.e.*, the battery is never discharged below a certain level:  $B(t)$  represents a scaling of the actual state of charge), hence undercharge may safely be neglected [6, Chapter 3].

#### D. Consumer Policies

At every time slot the consumer knows the maximum amount of energy  $x_t$  allowed by the DSO and responds by choosing an actual energy consumption  $z_t$ , using a *consumer policy*, which is defined as a mapping  $\pi : \vec{x} \rightarrow \vec{z}$ . More precisely, for any  $0 \leq s, t < T$ , we use the following definitions:

- Let  $\mathcal{X}_s(\vec{x}_{:t}) \subset (\mathbb{R}^+)^s$  be the set of sequences of feasible DSO control signals for the next  $s$  time steps, *i.e.*, that satisfy the service curve constraint (4) given (feasible) past signals  $\vec{x}_{:t} = (\dots, x_{-1}, x_0, \dots, x_t)$ . A “policy starting after time  $t$ ” is a mapping  $\mathcal{X}_{T-t}(\vec{x}_{:t}) \xrightarrow{\pi} (\mathbb{R}^+)^{T-t}$ .

- $\Pi_t(\vec{x}_{:t}, \vec{z}_{1:t}) \subset \{\pi : \mathcal{X}_{T-t}(\vec{x}_{:t}) \rightarrow (\mathbb{R}^+)^{T-t}\}$  is the set of all *valid* policies starting after time  $t$ , *i.e.* of all policies starting after time  $t$  such that the sequence of consumed energy decisions  $\vec{z}_{t+1:T}$  satisfies the constraints imposed by the service in (1) and the battery in (5). This set depends on the history  $(\vec{x}_{:t}, \vec{z}_{1:t})$ .
- A policy  $\pi \in \Pi_t(\vec{x}_{:t}, \vec{z}_{1:t})$  is called *causal* if for any  $\tau = t+1, \dots, T$  and for any  $\vec{y}, \vec{y}' \in \mathcal{X}_{T-t}(\vec{x}_{:t})$  such that  $\vec{y}_{t+1:\tau} = \vec{y}'_{t+1:\tau}$ , we have  $(\pi(\vec{y}))_s = (\pi(\vec{y}'))_s$ , for all  $s = t+1, \dots, \tau$ . In other words, the consumption decision taken by the policy at time  $\tau \in [t+1, \dots, T]$  depends only on the present and the past up to time  $\tau$ . We denote the set of all *causal* and *valid* policies starting after time  $t$  by  $\bar{\Pi}_t(\vec{x}_{:t}, \vec{z}_{1:t}) \subseteq \Pi_t(\vec{x}_{:t}, \vec{z}_{1:t})$ .

For example the *omniscient* policy  $\omega^*$ , introduced in (8) and which produces an optimal consumption sequence, is by definition a valid policy starting after time 0, but it is not causal (note that it is a hypothetical policy that we use only as a benchmark). The *greedy* policy, which consists in naively consuming as much as possible given the constraints of the service and of the battery, is a valid and causal policy.

One particular class of causal valid policies starting after time  $t$  is the class of *threshold policies*, which we now define. Given an adaptive threshold function  $h(s, b) \in [0, \bar{x}]$  that depends on the current time  $s > t$  and on past decisions  $\vec{z}_{1:s-1}$  via the current battery level  $b$ , the threshold policy  $\sigma_h$  is defined by

$$\begin{aligned} \sigma_h(\vec{x})(1) &\stackrel{\text{def}}{=} z_1 = \min\{x_1, h(0, B_0), f^{-1}(\bar{B})\}, \\ &\quad \text{if the policy starts at } t = 0 \\ \sigma_h(\vec{x})(s) &\stackrel{\text{def}}{=} z_s = \min\{x_t, h(s-1, B(\vec{z}_{1:s-1})), \\ &\quad f^{-1}(\bar{B} - B(\vec{z}_{1:s-1}))\}. \end{aligned} \quad (9)$$

where we denote the battery level at the end of time slot  $t$  by  $B(\vec{z}_{1:t}) = B_0 + \sum_{s=1}^t f(z_s)$ , given feasible decisions  $\vec{z}$ . In other words, this policy sets the consumption decision to the value of the threshold if this is possible given the constraints of the service and of the battery, else to the maximum allowed. By construction, a threshold policy starting after time  $t$  is valid and causal. The greedy policy is a threshold policy  $\sigma_{\bar{x}}$  (with  $h(s, b) = \bar{x}$  for all  $s, b$ ).

#### E. Definition of Optimal Policy

Intuitively, an optimal policy is a causal policy that minimizes the worst case cost for the consumer: it minimizes her energy consumption, hence aims to charge the battery at low power, while ensuring a full battery at the deadline, whenever the service reduction signals allow it. Such a policy is necessarily pessimistic: at any time step  $t$  it does not know the future, but using the knowledge of the service curve guarantees and of past reductions it protects the consumer against the worst future reduction signal (that maximizes her cost). We call this property *t-optimality*. The optimal policy updates its decisions at every time step, since observed service reductions may prove to be less constraining than the worst case. It is thus *t-optimal*, for all time steps  $t$ . We now define

rigorously the worst-case cost of a policy,  $t$ -optimality, and optimality.

For a valid policy starting after time  $t$ ,  $\pi \in \Pi_t(\vec{x}_{:t}, \vec{z}_{1:t})$ , the cost of  $\pi$  under control signals  $\vec{x}$  over  $t_1, \dots, t_2$  is denoted by  $c_{t_1:t_2}(\pi, \vec{x}) \stackrel{\text{def}}{=} C(\vec{z}_{t_1:t_2})$ , with  $\vec{z} = \pi(\vec{x})$  and  $C$  defined in (7). The worst case cost of  $\pi \in \Pi_t(\vec{x}_{:t}, \vec{z}_{1:t})$  is

$$\bar{c}(\pi) \stackrel{\text{def}}{=} \max_{\vec{y}_{t+1:T} \in \mathcal{X}_{T-t}(\vec{x}_{:t})} \{c_{t+1:T}(\pi, \vec{y}_{t+1:T})\}.$$

Informally, the worst case cost of a valid policy starting after  $t$  is the maximum cost that this policy may incur under all possible future control signals, given the constraints imposed by the past and present.

**Definition 1:** We say that a causal and valid policy starting at time  $t$ ,  $\pi^{*,t} \in \Pi_t(\vec{x}_{:t}, \vec{z}_{1:t})$ , is  $t$ -optimal if and only if  $\pi^{*,t}$  minimizes the worst-case cost over all other valid policies starting after time  $t$ .

Formally,  $\pi^{*,t}$  is  $t$ -optimal if and only if

$$\bar{c}(\pi^{*,t}) = \min_{\pi' \in \Pi_t(\vec{x}_{:t}, \vec{z}_{1:t})} \bar{c}(\pi').$$

Our risk-averse consumer wants to find a policy that is causal and  $t$ -optimal at every time  $t$ , taking into consideration what she can observe up to time  $t$ , and for every time  $t = 1, \dots, T$ . This motivates the following definitions.

**Definition 2:** We are given a causal and valid policy starting after  $t$ ,  $\pi \in \Pi_t(\vec{x}_{:t}, \vec{z}_{1:t})$ , some  $s > t$ , and some fixed feasible control signal  $\vec{y}_{t+1:s}$  given  $\vec{x}_{:t}$ . The restricted policy starting after time  $s$  given  $\vec{y}_{t+1:s}$ , denoted as  $\pi|_{\vec{y}_{t+1:s}}$ , is defined by

$$\pi|_{\vec{y}_{t+1:s}}(\vec{y}_{s+1:T}') \stackrel{\text{def}}{=} \vec{z}_{s+1:T}$$

where  $\vec{z}_{t+1:T} = \pi([\vec{y}_{t+1:s}, \vec{y}_{s+1:T}'])$  for any feasible signal  $\vec{y}_{s+1:T}' \in \mathcal{X}_{T-s}([\vec{x}_{:t}, \vec{y}_{t+1:s}])$ .

The restricted policy is well defined since, by causality, the decisions  $\vec{z}_{t+1:s}$  are the same irrespective of the future signal  $\vec{y}_{s+1:T}'$ . The restricted policy is valid and causal:  $\pi|_{\vec{y}_{t+1:s}} \in \Pi_s([\vec{x}_{:t}, \vec{y}_{t+1:s}], \vec{z}_{1:s})$ .

Informally, given a valid and causal policy  $\pi$  starting after time  $t$ , the restricted policy starting after time  $s > t$  is the policy obtained by applying  $\pi$  in the future after  $s$ , given that we now know the control signal  $\vec{x}_{t+1:s}$  up to time  $s$ .

**Definition 3:** We say that a causal and valid policy starting at time 0,  $\pi^* \in \Pi_0(\vec{x}_{:0})$ , is optimal if and only if for any time step  $0 \leq t < T$  and for any feasible control signal  $\vec{y}_{1:t} \in \mathcal{X}_t(\vec{x}_{:0})$ , the restricted policy  $\pi^*|_{\vec{y}_{1:t}}$  is  $t$ -optimal.

Informally, an optimal policy is a valid and causal policy that has the following property: for each time step  $t$ , after observing the  $t$  first control signals, the resulting restricted policy starting after  $t$  has a worst case cost that matches the smallest worst case cost of any valid policy (not necessarily causal) starting at  $t$ .

In the next section we prove by construction the existence of an optimal policy.

### III. EXISTENCE AND COMPUTATION OF AN OPTIMAL POLICY

In this section we present our main result: there exists a causal energy consumption policy for the consumer that

minimizes her cost in the worst-case scenario permitted by the service curve contract. We first show the main technical result, which is the construction of a  $t$ -optimal policy as a threshold policy. Then we show how to derive an optimal policy from any  $t$ -optimal policy.

#### A. Existence and Computation of a $t$ -Optimal Policy

The following theorem states that there exists a threshold policy that is  $t$ -optimal. This result is useful for determining an optimal policy.

**Theorem 1:** Consider an adjustable service specified using a load-switching service curve  $\beta$  with parameters  $t_0$ ,  $t_1$ , and  $\bar{x}$ , like in (3). Given past decisions  $\vec{z}_{1:t}$  and feasible past control signals  $\vec{x}_{:t}$  there exists a constant  $g_t$ , which can be computed as a function of past signals and decisions at the end of time slot  $t$ , such that the fixed threshold policy  $\sigma_{g_t}$  is  $t$ -optimal.

The constant  $g_t$  is given by  $g_t = h_T^*(t, B(\vec{z}_{1:t}))$ , with

$$h_T^*(s, b) = \begin{cases} f^{-1}\left(\frac{\bar{B}-b}{q_T(s)+1}\right), & \text{if } \frac{\bar{B}-b}{q_T(s)+1} \leq f(r_T(s)), \\ f^{-1}\left(\frac{\bar{B}-b-f(r_T(s))}{q_T(s)}\right), & \text{if } \frac{\bar{B}-b}{q_T(s)+1} > f(r_T(s)) \\ & \text{and } q_T(s) > 0, \\ \infty, & \text{otherwise,} \end{cases} \quad (10)$$

where  $q_T(s) \in \mathbf{N}$  and  $0 \leq r_T(s) < \bar{x}$  are the unique values such that  $\max(H_T(s), 0) = q_T(s)\bar{x} + r_T(s)$ , with

$$H_T(s) \stackrel{\text{def}}{=} \beta(t_1) - \sum_{t=T-t_1+1}^s x_t. \quad (11)$$

The proof of this result is lengthy and relies on a game-theoretic approach. It is given in Section IV.

#### B. Existence and Computation of an Optimal Policy

The following theorem constructs an optimal policy from any  $t$ -optimal policy and can thus be combined with the previous section to obtain an optimal policy.

**Theorem 2:** Assume that we have a method to compute a  $t$ -optimal policy  $\pi^{*,t} \in \Pi_t(\vec{x}_{:t}, \vec{z}_{1:t})$  at every time  $t = 0, 1, \dots, T$  from any past control signals and decisions  $(\vec{x}_{:t}, \vec{z}_{1:t})$  (note that  $\pi^{*,t}$  must be feasible and causal).

Let  $\vec{x}_{:0}$  be some arbitrary feasible sequence of control signals. For  $t = 1, \dots, T$ , define the consumption  $z_t^*$  iteratively as follows:

- $z_1^*$  is the decision taken by the policy  $\pi^{*,0}$  given  $\vec{x}_{:0}$
- $z_t^*$  is the decision taken by the policy  $\pi^{*,t-1}$  given  $(\vec{x}_{:(t-1)}, \vec{z}_{1:(t-1)}^*)$

Then the policy  $\pi^*$  defined by  $(x_1, \dots, x_T) \mapsto (z_1^*, \dots, z_T^*)$  is optimal.

The proof is given in Section V. Note that the policy  $\pi^{*,t}$  in the theorem is a feasible and causal policy starting after time  $t$  whereas the policy  $\pi^*$  is a feasible and causal policy starting after time 0.

A direct consequence of Theorem 2 is the following:

**Corollary 1:** The threshold policy  $\sigma_{h_T^*}$ , where  $h_T^*$  is defined in (10), is an optimal policy.

Thus, Algorithm 1 implements an optimal policy.

---

**Algorithm 1** Optimal (Causal) Policy
 

---

**Require:**  $\vec{x}_{:0}$ ,  $B_0$ ,  $\bar{B}$ ,  $T$ , load-switching service curve  $\beta$  with parameters  $t_0$ ,  $t_1$ ,  $\bar{x}$ , like in (3)

**Ensure:**  $\vec{x}_{:0}$  satisfies service curve  $\beta$

```

1:  $H \leftarrow \beta(t_1) - \sum_{s=T-t_1+1}^0 x_s$ 
2: Current battery level is  $B \leftarrow B_0$ 
3: for  $t = 1$  to  $T$  do
4:    $q \leftarrow \lfloor \max\{0, H\} / \bar{x} \rfloor$ 
5:    $r \leftarrow \max\{0, H\} - q\bar{x}$ 
6:    $h \leftarrow \bar{x}$ 
7:   if  $\bar{B} - B / q + 1 \leq f(r)$  then
8:      $h \leftarrow f^{-1}((\bar{B} - B) / (q + 1))$ 
9:   else if  $q > 0$  then
10:     $h \leftarrow f^{-1}([\bar{B} - B - f(r)] / q)$ 
11:   end if
12: Receive signal  $x_t$ 
13: Consume  $z_t \stackrel{\text{def}}{=} \min\{x_t, h\}$ 
14:  $B \leftarrow B + f(z_t)$ 
15:  $H \leftarrow H - x_t$ 
16: end for

```

---

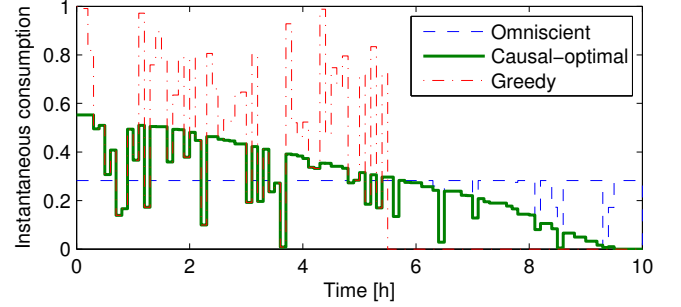
### C. Numerical Illustration

We compare numerically the cost of using Algorithm 1 to the costs of the omniscient policy  $\omega^*$  and of the greedy policy  $\gamma \equiv \sigma_{\bar{x}}$  that consists of charging the battery by consuming the maximum allowed energy in each time slot. We consider that the increase in battery level is given by  $f$  defined in (6) with a charging efficiency  $\eta = 90\%$  and an internal resistance  $\rho$  such that the resistive loss when charged at maximum intensity is  $1 - \alpha = 30\%$  (hence  $\rho = \frac{\eta\alpha}{\bar{x}}$ ).

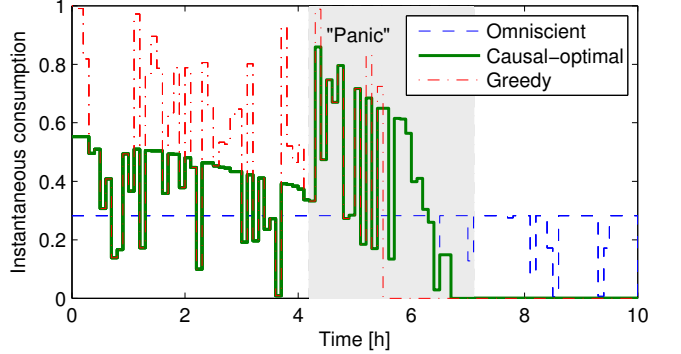
We express the amount of required charge  $\bar{B} - B_0$  as the fraction  $w$  of the time before the deadline that the greedy policy  $\gamma$  needs to charge the battery when the control signal always allows maximum energy per time slot  $\bar{x}$ . Specifically,  $\bar{B} - B_0 = wT \cdot f(\bar{x})$ . Thus, a small  $w$  indicates that there is a lot of flexibility in the schedule.

We consider a load-switching curve  $\beta$  with periodicity of the derivative of  $t_1 = 24\text{h}$ , maximum service interruption time of  $t_0 = 7\text{h}$ , and maximum power  $\bar{x}$ . The battery needs to be charged within  $T = 10\text{h}$  from the present time. We generate random control signals  $\vec{x}_{(2-t_1):T}$  (i.e., past signals  $\vec{x}_{(2-t_1):0}$  and signals during the consumption period  $x_{1:T}$ ) that satisfy the service curve constraint: all signals previous to  $x_{2-t_1}$  are set to  $\bar{x}$ , and for each  $t = 2 - t_1, \dots, T$ , we pick  $x_t$  uniformly at random between the least allowed service  $\max\{0, \beta(t_1) - \sum_{s=t-t_1+1}^{t-1} x_s\}$  and the maximum service  $\bar{x}$ . For such randomly generated control signals (that by definition satisfy the service curve) we evaluate the cost of the three policies.

First, for a given control signal, in Figure 2 we show a run of the three policies in two scenarios: when the deadline  $T$  remains the same throughout the entire period (Figure 2(a)), and when it is changed mid-charge by an anxious consumer (Figure 2(b)). In both cases, the consumption curve of the greedy policy coincides with the control signal (maximum power allowed), since it consumes as much as it is given (until



(a) Deadline is fixed  $T = 10\text{h}$



(b) When the shaded area begins, the deadline is advanced from  $T = 10\text{h}$  to the end of the shaded area,  $T \approx 7\text{h}$

Fig. 2. Sample runs when the battery can be filled by consuming  $\bar{x}$  per time slot for a fraction  $w = 1/3$  of the time before the deadline. We show two cases: 2(a) – when the consumer allows the causal optimal policy to run until the end, and 2(b) – when the consumer panics and advances the deadline mid-charge. The greedy policy finishes charging first, but consumes the most. The threshold of the omniscient policy is visible, as well as the evolving (decreasing) threshold of the causal optimal policy.

it charges the battery). The adjustable threshold followed by the optimal policy is visible: the optimal policy consumes less than the threshold only when it is not allowed to reach it by the control signals. In the second case, we see that the optimal policy adapts to the new, nearer deadline and succeeds to charge the battery before reaching it, while still consuming less than the greedy policy in period in which the two overlap.

Note that if the battery cannot be filled before the deadline, the three policies behave identically and consume everything that is allowed by the control signal. Indeed, the omniscient policy knows the future, the causal policy considers the worst case future signal, and the greedy policy does this by default. Thus, the three policies can only be distinguished when the control signals allow the battery to be filled before the deadline, which all three achieve (by construction). In this case, the amount of consumed energy differs for the three. For this reason, in what follows we only consider cases where the battery is filled by the three policies and compare their resulting energy consumption.

Consider a given required charge  $w = \frac{\bar{B} - B_0}{Tf(\bar{x})}$ , a given control signal  $\vec{x}$  that allows a full charge of the battery, and a policy  $\pi$  that takes advantage of  $\vec{x}$  and sets a consumption sequence  $\vec{z} = \pi(\vec{x})$  that indeed achieves a full battery  $B(\pi(\vec{x})) = \bar{B}$  at  $T$ . We define the relative cost of  $\pi$  under signal  $\vec{x}$  as the ratio between its total cost  $\sum_{t=1}^T pz_t =$



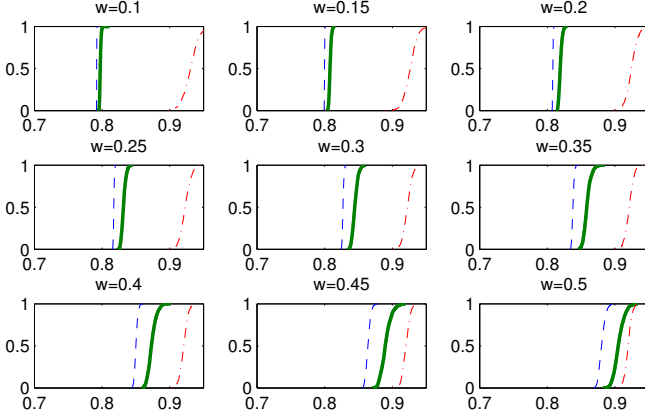


Fig. 3. Empirical CDF of the relative cost (12) for the three policies for 1000 randomly generated control signals that satisfy the load-switching service curve; the performance of the omniscient policy is represented in dashed lines, the causal-optimal policy in full lines, and the greedy policy in dash-dot lines. Each plot corresponds to a different required charge  $w$  that can be obtained by consuming the maximum allowed energy per slot  $\bar{x}$  during  $wT$  slots.

$c_{1:T}(\pi, \vec{x}) + K\bar{B}$  and the cost in a hypothetical setting when the greedy policy is used without any restrictions from the DSO ( $x_t = \bar{x}$  for all  $t$ ). The numerator is in this case just the energy bill, since  $\pi$  fills the battery and there is no final penalty. The value of the denominator is the cost when the battery is charged by consuming the maximum energy per time slot (in the least efficient way). Specifically, the relative cost is

$$c_r(\pi, \vec{x}) = \frac{c_{1:T}(\pi, \vec{x}) + K\bar{B}}{pwT\bar{x}} = \frac{\sum_{t=1}^T z_t}{wT\bar{x}} \in [0, 1]. \quad (12)$$

Finally, we can thus compare the performance of the three policies in terms of the relative cost for different values of  $w$ . In Figure 3, for 9 values of  $w$  we plot the CDF of the relative cost for the three policies obtained over 1000 random control signals that satisfy the service curve. We observe that the cost of the causal optimal policy is typically closer to the one of the omniscient policy. We stress that for any control signal  $\vec{x}$  the cost of the three policies is always ordered:

$$c(\omega^*, \vec{x}) \leq c(\pi^*, \vec{x}) \leq c(\gamma, \vec{x}),$$

i.e.,  $\pi^*$  is always better than  $\gamma$ .

#### IV. PROOF OF THEOREM 1

Let us consider that the Distribution System Operator (DSO) is adversarial and wants to maximize the cost of the consumer. We study this scenario because the consumer wants to find the best response against the worst control signal allowed by the service curve. In reality the DSO will most likely take milder actions that suit such objectives as balancing energy flows, or controlling network state.

For ease of presentation, we prove the theorem for  $t=0$  (hence no past decisions), with an initial battery level  $B_0$ , without loss of generality. Indeed, the past decisions contribute only to increasing the initial level of the battery. Let us proceed with the steps needed to prove Theorem 1.

#### A. Game-Theoretical Approach

Consider the (virtual) game between a player 1, called “consumer”, and a player 2, called “DSO”. The consumer chooses a policy  $\pi \in \Pi_0(\vec{x}_{:0})$  and DSO chooses a control signal  $\vec{x} \in \mathcal{X}_T(\vec{x}_{:0})$ . The benefit to the consumer is  $-c(\pi, \vec{x})$ , to DSO it is  $c(\pi, \vec{x})$ . This is the simplest of all games, namely a zero-sum simultaneous game with 2 players. Assume that best responses are well defined and let  $x^*(\pi)$  [resp.  $\pi^*(\vec{x})$ ] denote the set of best responses to  $\pi$  [resp. to  $\vec{x}$ ].

We show that 0-optimal policies correspond to a Nash Equilibrium [5], [11] of this game; more precisely, a consumer’s 0-optimal policy can be obtained as her best response in a Nash equilibrium:

**Lemma 1:** If  $(\pi, \vec{x})$  is a Nash equilibrium, then  $\pi$  is 0-optimal.

*Proof:* Since  $\vec{x}$  is a best response to  $\pi$ , then  $c(\pi, \vec{x}') \leq c(\pi, \vec{x})$  for any  $\vec{x}' \in \mathcal{X}_T(\vec{x}_{:0})$ . Thus  $\bar{c}(\pi) = c(\pi, \vec{x})$ .

Similarly  $\pi$  is a best response to  $\vec{x}$ , thus  $c(\pi, \vec{x}) \leq c(\pi', \vec{x})$  for any  $\pi' \in \Pi_0(\vec{x}_{:0})$ . Now  $c(\pi', \vec{x}) \leq \bar{c}(\pi')$ , by definition of a max, and hence  $\bar{c}(\pi) = c(\pi, \vec{x}) \leq c(\pi', \vec{x}) \leq \bar{c}(\pi')$ , which shows that  $\pi$  is 0-optimal. ■

(Note that it is not clear that the converse is true.)

Therefore the problem is to find a Nash equilibrium for this game, if it exists. In other words, the “worst case”  $\vec{x}^*$  that we are looking for is given by a Nash equilibrium. We can find it as follows: first, we determine the best response of the consumer to a given DSO control signal, i.e., the structure of the omniscient policy  $\omega^*(\vec{x})$ ; second, we characterize the DSO’s best response to this consumer policy:  $\vec{x}^*(\omega^*)$ ; finally, we aim to solve the fixed point equation  $\vec{y} = \vec{x}^*(\omega^*(\vec{y}))$ .

#### B. The Omniscient Policy. Best Consumer Response

In the following lemma we determine the structure of the omniscient policy  $\omega^*(\vec{x})$  which constitutes the best response of the consumer to a given (known) control signal  $\vec{x}$ . We prove it in Appendix A.

**Lemma 2:** Consider a battery with an initial charge  $B_0$ , and a sequence of control signals  $\vec{x}_{:0}$ . The omniscient cost-minimizing policy  $\omega^* \in \Pi_0(\vec{x}_{:0})$  defined in (8) has the following structure: there exists a value  $0 < h^*(\vec{x}) < \bar{x}$  such that  $\omega^*(\vec{x})(t) = \min\{x_t, h^*(\vec{x})\}$ .

Thus, if there exists a Nash equilibrium, then there exists a “worst case” signal  $\vec{x}^*$  for the consumer. Assume that we can compute this  $\vec{x}^*$ . Lemma 2 characterizes the best response of the consumer to this specific signal. Then a natural candidate for a causal policy is the fixed threshold policy  $\sigma_{h^*(\vec{x}^*)}$ . Indeed, for the control signal  $\vec{x}^*$ , this choice minimizes the cost of the consumer. The following questions remain unanswered at this point.

- 1) What is the “worst-case” signal  $\vec{x}^*(\sigma_h)$  for a consumer implementing a fixed threshold policy  $\sigma_h$  ( $h > 0$ )?
- 2) Does there exist a Nash equilibrium?

We address them in the next section.

#### C. The Best Response of the DSO. Nash Equilibrium

In the following lemma we answer the first question: given that the consumer implements a fixed threshold policy  $\sigma_h$ ,

what is the worst case signal for the consumer? We prove it in Appendix B.

*Lemma 3:* Consider that the consumer with an initial battery level  $B_0$  to be filled by time  $T$  implements a fixed threshold policy  $\sigma_h$  defined in (9) with  $h > 0$  constant. Given that the consumer has received past signals  $\vec{x}_{:0} = (\dots, x_{-1}, x_0)$  (that themselves satisfy the service curve constraints), an adversarial DSO that wants to maximize the cost of the consumer, subject to the  $\beta(\cdot)$  service curve constraints, should send the following signal ( $r_T(0)$  is defined in Theorem 1):

$$\begin{aligned} x_1^* &= \dots = x_T^* = \bar{x}; \\ x_{T+1}^* &= r_T(0); \\ x_{T+2}^* &= \dots = x_T^* = 0. \end{aligned}$$

Intuitively, if the DSO delivers maximum service in the beginning, it alleviates her service constraints for the end of the desired charging period. This allows her to interrupt the service after some time until the deadline  $T$ . We stress again that this malicious behaviour may be encountered in practice only coincidentally, specifically if it happens to suit the real interests of the DSO. We study this scenario because the consumer wants to be shielded from interruptions.

It is important to note that for *any* fixed threshold policy  $\sigma_h$ , the signal that maximizes cost is the  $\vec{x}^*$  in Lemma 3 and *does not depend on  $h$* . Moreover, it can be computed at the beginning of time step 1.

We can now address the second question, that concerns the existence of a Nash equilibrium. We claim that  $(\sigma_{h^*}(\vec{x}^*), \vec{x}^*)$  is a Nash equilibrium. Indeed, Lemma 2 shows that  $c(\sigma_{h^*}(\vec{x}^*), \vec{x}^*) \leq c(\pi, \vec{x}^*)$ , for all  $\pi \in \Pi(\vec{x}_{:0}, B_0)$ , i.e., the consumer has no incentive to choose a different policy. Then, Lemma 3 shows that  $c(\sigma_{h^*}(\vec{x}^*), \vec{x}^*) \geq c(\sigma_{h^*}(\vec{x}^*), \vec{x})$ , for all  $\vec{x} \in \mathcal{X}_T(\vec{x}_{:0})$ . Thus,  $\sigma_{h^*}(\vec{x}^*)$  is a 0-optimal policy. Since  $\vec{x}^*$  can be computed at the beginning of time step 1,  $\sigma_{h^*}(\vec{x}^*)$  is causal. This concludes the proof.

## V. PROOF OF THEOREM 2

We dedicate this short section to the proof of Theorem 2 because of its generality.

By construction,  $\pi^*$  is causal. Let us show that, for any  $t$  and any  $\vec{x}_{1:t} \in \mathcal{X}_t(x_{:0})$ , the policy  $\pi^*|_{\vec{x}_{1:t}}$  is  $t$ -optimal. We use backwards induction on  $t$ . In the base case when  $t = T - 1$  we are given all past signals  $\vec{x}_{1:T-1}$  and decisions  $\vec{z}_{1:T}$ . Thus, by definition,  $\pi^*|_{\vec{x}_{1:T-1}} = \pi^{*,T-1} \in \bar{\Pi}_{T-1}(\vec{x}_{:T-1}, \vec{z}_{1:T-1})$  which is by hypothesis  $(T - 1)$ -optimal.

Suppose now that  $\pi^*|_{\vec{x}_{1:t}}$  is  $t$ -optimal for some  $t > 0$  and any  $\vec{x}_{1:t} \in \mathcal{X}_t(\vec{x}_{:0})$ . Then for any  $\vec{x}_{1:t-1} \in \mathcal{X}_{t-1}(\vec{x}_{:0})$ :

$$\begin{aligned} \bar{c}(\pi^*|_{\vec{x}_{1:t-1}}) &= \max_{\vec{y}_{t:T} \in \mathcal{X}_{T-t+1}(\vec{x}_{t-1})} c_{t:T}(\pi^*|_{\vec{x}_{1:t-1}}, \vec{y}_{t:T}) \\ &= \max_{\vec{y}_{t:T} \in \mathcal{X}_{T-t+1}(\vec{x}_{t-1})} \{ \phi(\pi^*|_{\vec{x}_{1:t-1}}(y_t)) + \\ &\quad c_{t+1:T}(\pi^*|_{[\vec{x}_{1:t-1} y_t]}, \vec{y}_{t+1:T}) \} = \max_{y_t} \{ \phi(\pi^*|_{\vec{x}_{1:t-1}}(y_t)) \\ &\quad + \max_{\vec{y}_{t+1:T} \in \mathcal{X}_{T-t}([\vec{x}_{1:t-1} y_t])} c_{t+1:T}(\pi^*|_{[\vec{x}_{1:t-1} y_t]}, \vec{y}_{t+1:T}) \} \\ &= \max_{y_t} \{ \phi(\pi^{*,t-1}(y_t)) + \bar{c}(\pi^*|_{[\vec{x}_{1:t-1} y_t]}) \} \\ &\leq \max_{y_t} \{ \phi(\pi^{*,t-1}(y_t)) + \bar{c}(\pi^{*,t-1}|_{y_t}) \} = \bar{c}(\pi^{*,t-1}). \end{aligned}$$

We have used the fact that all referenced policies are causal. The inequality holds thanks to the induction hypothesis: the policy  $\pi^*|_{[\vec{x}_{1:t-1} y_t]}$  is  $t$ -optimal, hence its worst-case cost is lower than the worst-case cost of any other policy, in particular that of the restricted policy  $\pi^{*,t-1}|_{y_t}$ .

We have shown that  $\pi^*|_{\vec{x}_{1:t}}$  is  $t$ -optimal, since its worst-case cost is lower than that of the  $t$ -optimal  $\pi^{*,t-1}$ .

Incidentally, we have shown that the causal-optimal policy in Algorithm 1 corresponds to a subgame-perfect Nash equilibrium [5] of the multi-stage game between the consumer and the DSO.

## VI. CONCLUSION

We have characterized a causal (optimal) charging policy that minimizes at each time step the worst-case cost for a non-ideal battery that is subjected to service reductions with a service curve guarantee. This policy takes the form of an adaptive threshold policy. The cost comprises the electricity consumption and the disutility of having a discharged battery at the deadline. The type of cost function we consider is tailored such that the battery is charged at the deadline unless the overall allowed consumption is insufficient. Thus the service reductions are rendered as transparent as possible for the consumer. Numerical evaluations indicate that the cost of the optimal causal policy tends to be close to the one of the omniscient policy.

## APPENDIX A PROOF OF LEMMA 2

*Proof:* We want to minimize  $\sum_{s=1}^T p z_s - K f(z_s)$  over  $z_s \leq x_s$ , such that  $\sum_{s=1}^T f(z_s) \leq \bar{B} - B_0$ . We write the Lagrangian [3] of this convex optimization problem

$$\mathcal{L} = \sum_{s=1}^T p z_s - K f(z_s) + \sum_{s=1}^T \mu_s (z_s - x_s) + \lambda \left[ B_0 - \bar{B} + \sum_{s=1}^T f(z_s) \right].$$

Then writing the KKT conditions we get:

$$\begin{aligned} z_s &= f'^{-1} \left( \frac{p + \mu_s}{K - \lambda} \right); \quad \mu_s [z_s - x_s] = 0, \quad \mu_s \geq 0; \\ \lambda \left[ B(0) - \bar{B} + \sum_{s=1}^T f(z_s) \right] &= 0, \quad \lambda \geq 0. \end{aligned}$$

We distinguish the inactive and active constraints: If for some  $s$  we have  $z_s = x_s$ , then we find the value of the multiplier  $\mu_s = (K - \lambda) f'(x_s) - p > 0$  and, by its positivity, it follows that  $x_s < h(\lambda) \stackrel{\text{def}}{=} f'^{-1} \left( \frac{p}{K - \lambda} \right)$ . Otherwise, if  $z_s < x_s$ , then  $z_s = h(\lambda) \leq x_s$ .

We have found that a threshold  $h = h(\lambda)$  exists for which the stated result holds. We now determine this threshold.

Assume that the overflow constraint is inactive (i.e., the battery is not filled by the deadline). Then we have  $\lambda = 0$ , and  $h(0) = f'^{-1} \left( \frac{p}{K} \right)$ . But  $K$  was chosen such that  $\frac{K}{p} > \frac{1}{f'(\bar{x})}$ , and hence  $h(0) > \bar{x} \geq x_s$  for all  $s$ . Thus, since  $K$  is large, the optimal solution is to always charge at maximum intensity  $z_s = x_s$ , since all  $x_s$  are less than the threshold  $h(0)$ .



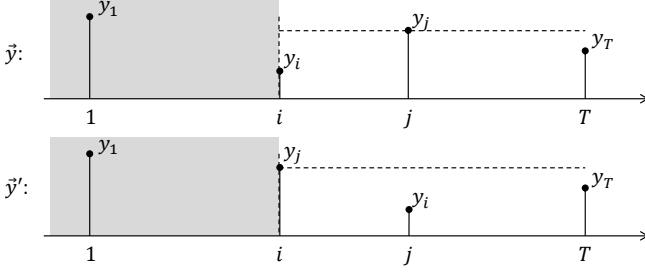


Fig. 4. The “flip” operation done at  $i$ . The control sequence  $\bar{y}'$  is obtained from  $\bar{y}$  by flipping  $y_i$  and  $y_j$ , where  $y_j$  is the largest among  $y_i, \dots, y_T$ .

Finally, let the overflow constraint be active (i.e., the battery is filled by the deadline). We must determine  $h(\lambda)$  such that

$$B_0 - \bar{B} + \sum_{s: x_s < h(\lambda)} f(x_s) + \sum_s \mathbf{1}_{\{x_s \geq h(\lambda)\}} f(h(\lambda)) = 0.$$

Consider wlog. that the  $x_s$  are sorted in decreasing order. Due to the monotonicity of  $f$ , the threshold  $h(\lambda)$  is unique. Consider  $t^*$ , the largest  $t$  such that  $x_t \geq h(\lambda) > x_{t+1}$  (by convention  $x_{T+1} = 0$ ). In other words,  $t^*$  satisfies  $\bar{B} - B_0 - \sum_{s=t^*+1}^T f(x_s) \leq t^* f(x_{t^*})$ . Hence,

$$h(\lambda) = f^{-1} \left( \frac{\bar{B} - B_0 - \sum_{s=t^*+1}^T f(x_s)}{t^*} \right).$$

*Note:* If we consider  $f(z) = \eta z - \rho z^2$ , then  $t^*$  is the largest  $t$  such that

$$t\eta^2 + 4\rho \left( B_0 + \sum_{s=t+1}^T f(s) - \bar{B} \right) < 0.$$

## APPENDIX B

### PROOF OF LEMMA 3

*Proof:* Denote  $\sigma(t) = \sum_{s=t-t_1+1}^0 x_s$ . Since  $x_s \geq 0$ , we have that  $(\sigma(t))_{t=1}^T$  is a decreasing sequence. From Theorem 2 of [7] it follows that there are  $T$  linear constraints  $(\mathcal{C}_1), \dots, (\mathcal{C}_T)$  imposed by the service curve  $\beta$  on the control signals, where  $(\mathcal{C}_t)$  is:

$$(\mathcal{C}_t): \quad \sigma(t) + \sum_{s=1}^t x_s \geq \beta(t_1).$$

Suppose we have a sequence  $\bar{x}^*$  of signals like in the lemma statement. By definition they satisfy  $(\mathcal{C}_T)$  (with equality if  $H > 0$ , or otherwise  $\bar{x}^* \equiv 0$ ). Since  $x_s^* = 0$  for all  $s \geq \tau + 2$ , and since  $\sigma(t)$  is a decreasing sequence, the  $\bar{x}^*$  also satisfy  $(\mathcal{C}_{\tau+1}), \dots, (\mathcal{C}_{T-1})$ . Moreover, since the  $x_s^*$  take the maximum value  $x_s^* = \bar{x}$  for all  $1 \leq s \leq \tau$ , and since  $\bar{x}_0$  is a valid sequence of controls, it follows that the remaining constraints  $(\mathcal{C}_1), \dots, (\mathcal{C}_\tau)$  are also satisfied.

Consider a sequence of signals  $\bar{y} = (y_1, \dots, y_T)$  that satisfies the service curve constraints  $(\mathcal{C}_t)$ . Pick a time instant  $i$  such that there exists a further time instant  $i' > i$  with  $y_i < y_{i'}$ . Denote  $j = \arg \max\{y_{i'} : i' > i\}$ . Then the “flipped” sequence at  $i$  defined as  $\bar{y}' = (y_1, \dots, y_j, \dots, y_i, \dots, y_T)$  (see

Figure 4) also satisfies the service curve constraint and, under a consumer threshold policy  $\sigma_h$ , has a cost that is at least as large as the one of the original control sequence. The first claim is true: only constraints  $(\mathcal{C}_i), \dots, (\mathcal{C}_j)$  are impacted by the flip, but since  $y_j > y_i$  we have that

$$\sum_{s=1}^t y'_s \geq \sum_{s=1}^t y_s \geq \beta(t_1) - \sigma(t)$$

for all  $i \leq t \leq j$  (since  $\bar{y}$  satisfies the constraints). To show the second claim, consider the cost difference between the two scenarios

$$c(\sigma_h, \bar{y}) - c(\sigma_h, \bar{y}') = \sum_{s=i}^T \phi(z_s) - \phi(z'_s),$$

where  $\bar{z}$  and  $\bar{z}'$  are the sequences of consumed energy at each time step by a threshold policy under signals  $\bar{y}$  and  $\bar{y}'$ , respectively.

Notice that we have that  $\sum_s f(z_s) = \sum_s f(z'_s)$ . Indeed, if for one signal sequence the battery is filled, then it is also filled for the other one, and, vice-versa, if the battery is not filled for one, then it is not filled for the other. In the latter case the same energy is consumed (only the order differs). Thus a difference in cost may only occur when the battery is filled before the deadline, and it originates from the linear consumption term in  $\phi$ . In this setting, consider the time slot  $k$  when the battery reaches the maximum charge for signal  $\bar{y}$  and the corresponding time slot  $k'$  for  $\bar{y}'$ .

If  $k \leq i$ , the cost difference is zero (at time  $i$  the consumption remains the same, despite a larger allowed energy).

Suppose that  $z_i = z'_i$ . Then either  $h < \min\{y_i, y'_i\}$ , or  $k' = k = i$  and the required energy to finalize the battery charge is less than  $y_i$ . In both cases the flip does not impact the cost.

Otherwise, we have  $z_i < z'_i$ . It must be that  $i \leq k' \leq k$ . Suppose  $k' = i < k$ . We have

$$\begin{aligned} c(\sigma_h, \bar{y}) - c(\sigma_h, \bar{y}') &= \sum_{s=i}^k \phi(z_s) - \phi(z'_s) \\ &= \sum_{s=i}^k \phi(z_s) - \phi(z'_i) = p \left( \sum_{s=i}^k z_s - z'_i \right) < 0, \end{aligned}$$

since  $f(z'_i) = \sum_{s=i}^k f(z_s) \geq f\left(\sum_{s=i}^k z_s\right)$  and  $f' > 0$ .

Otherwise,  $k' > i$ . Since consumption at times steps  $s = i+1, \dots, k'-1$  coincides ( $z_s = z'_s$ ), we have that

$$c(\sigma_h, \bar{y}) - c(\sigma_h, \bar{y}') = p \sum_{s=k'+1}^k (z_s + z_i + z_{k'} - z'_i - z'_{k'}).$$

Necessarily  $z_{k'} \geq z'_{k'}$ . Since  $y'_i > y'_s$  for all  $s \geq i$  (by definition), it follows that  $z'_i > z'_{k'}$  and  $z'_i > z_s$  for all  $s \geq i$ . We apply Lemma 4 (with  $y := z'_i$ ,  $x_0 := z'_{k'}$ ) and conclude that the difference in cost is negative.

Hence for any signal that satisfies the service curve we can apply the “flip” operation until the resulting signal is sorted decreasingly, and the cost will not decrease. Thus, a cost-maximizing signal can be rendered decreasing.

Consider now a decreasing signal  $\bar{y}$ , and let us increase the allowed consumption in the beginning of the period and

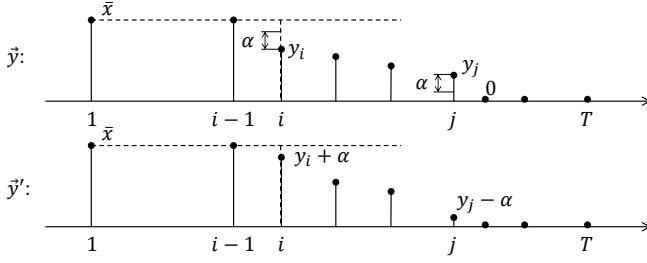


Fig. 5. The “ $\alpha$ -unbalance” operation on a decreasing sequence  $\bar{y}$ . The control sequence  $\bar{y}'$  is obtained from  $\bar{y}$  by adding  $\alpha$  to  $y_i$  and by subtracting  $\alpha$  from  $y_j$ , where  $y_i$  is the largest strictly less than  $\bar{x}$ , and  $y_j$  is the smallest strictly larger than 0.

decrease it toward the end. Formally, let  $i = \min\{\ell : y_\ell < \bar{x}\}$ ,  $j = \max\{\ell : y_\ell > 0\}$ , and  $\alpha = \min\{\bar{x} - y_i, y_j\}$ . The “ $\alpha$ -unbalanced” signal defined as  $\bar{y}' = (\bar{x}, \dots, y_i + \alpha, \dots, y_j - \alpha, 0, \dots, 0)$  (see Figure 5) satisfies the service curve constraint and increases the cost for the consumer. The first claim is readily verified by checking  $(\mathcal{C}_k), \dots, (\mathcal{C}_\ell)$ .

For the second claim, if  $h \leq y_i$  then using  $\bar{y}'$  the cost is either the same (if  $z_j < y_j - \alpha$ ), or it is strictly larger, as we are limiting consumption at time  $j$  while the same energy is consumed at time  $i$  in the two scenarios.

If  $h > y_i$ , then we distinguish two cases:

1) The battery is not filled using  $\bar{y}$ . We have

$$f(\min\{y_i + \alpha, h\}) + f(y_j - \alpha) \leq f(y_i + \alpha) + f(y_j - \alpha) \leq f(y_i) + f(y_j),$$

since  $f$  is concave increasing and  $y_i \geq y_j$ . Thus, the battery is not filled using  $\bar{y}'$  either. The cost increases, since  $\phi$  is convex decreasing and

$$\phi(\min\{y_i + \alpha, h\}) + \phi(y_j - \alpha) \geq \phi(y_i + \alpha) + \phi(y_j - \alpha) \geq \phi(y_i) + \phi(y_j) = \phi(z_i) + \phi(z_j).$$

2) The battery is filled using  $\bar{y}$ . Consider the time slot  $k$  when this occurs. Since  $y'_i > y_i$  and  $y'_s = y_s$  for all  $i < s < j$ , it follows that at the beginning of time slot  $k$  a larger battery level is attained by using  $\bar{y}'$  than by using  $\bar{y}$ . Thus, if the battery is filled for  $\bar{y}'$ , then  $f(z'_i) + f(z'_k) = f(z_i) + f(z_k)$ , where  $z'_i > z_i = y_i$  and  $z'_k < z_k \leq y_k \leq y_i$ , and hence the cost increases (by a similar argument to the “flip” case). Otherwise, necessarily  $k = j$  and  $y_j - \alpha < z_k$ . Thus,

$$\phi(\min\{y_i + \alpha, h\}) + \phi(z'_k) \geq \phi(y_i + \alpha) + \phi(y_j - \alpha) \geq \phi(z_i) + \phi(z_k).$$

Hence, for any decreasing signal we can apply successively the “unbalancing” operation until the first  $\ell$  components take the maximum value  $\bar{x}$ , the  $\ell+1$ -st component has some positive value, and the last  $T-\ell-1$  components are zero.

Let us now prune such an “unbalanced” signal  $\bar{y}$ . Let  $i = \max\{\ell : y_\ell > 0\}$  and define  $\alpha = \min\{y_i, \beta(t_1) - \sigma(T) - \sum_j y_j\}$ . The “pruned” signal is  $\bar{y}' = (\bar{x}, \dots, y_i - \alpha, 0, \dots, 0)$ . Then  $\bar{y}'$  satisfies the service curve constraints and has a higher consumer cost than  $\bar{y}$  for a threshold policy  $h$ .

Indeed, the first claim is verified since  $(\mathcal{C}_i), \dots, (\mathcal{C}_T)$  are verified by construction, and, due to the fact that  $y_1 = \dots = y_{i-1} = \bar{x}$  so are the rest of the constraints. The second

claim is also verified, since consumption is reduced, and  $\phi$  is decreasing.

Thus, consider the signal  $\bar{y}^*$  that maximizes cost for a threshold policy  $h$ . It follows that it can be transformed using the three operations above to the signal  $\bar{x}^*$  in the statement. ■

**Lemma 4:** If  $f : [0, b] \rightarrow \mathbf{R}_+$  is a continuously differentiable concave increasing function with  $b > 0$ , and the values  $y, x_0, x_1, \dots, x_n$ ,  $n \geq 2$  are such that: (i) there exists  $1 \leq k \leq n$  such that  $x_0 \leq x_k$ , (ii)  $y > \max_i x_i$ , and (iii)  $f(y) + f(x_0) = \sum_{i=1}^n f(x_i)$ , then we have that  $y + x_0 \geq \sum_{i=1}^n x_i$ .

*Proof:* For  $n = 2$ , assume wlog that  $x_0 \leq x_1$ . If  $x_0 > x_2$ , then  $f(x_0) + f(y) > f(x_2) + f(x_1)$ , which contradicts (iii). Hence, it is also true that  $x_0 \leq x_2$  and consider  $x_1 \leq x_2$ , wlog. We have that  $f(y) - f(x_2) = f(x_1) - f(x_0)$ . If  $f(x_1) - f(x_0) = 0$ , then the statement is trivially true. Otherwise, there exist  $u$  and  $v$ , ( $x_0 \leq u \leq x_1$  and  $x_2 \leq v \leq y$ ) such that

$$f(x_1) - f(x_0) = f'(u)(x_1 - x_0) \text{ and } f(y) - f(x_2) = f'(v)(y - x_2).$$

Since  $f$  is increasing and concave, it follows that  $f'(u) > f'(v) > 0$ , and thus that  $x_1 - x_0 < y - x_2$ .

For  $n > 2$ , we have that  $\sum_{i \neq k} x_i \leq y$ . Otherwise  $f(y) < \sum_{i \neq k} f(x_i)$ , which again leads to a violation of (iii). Denote

$$x_{-k} \stackrel{\text{def}}{=} \sum_{i \neq k} x_i.$$

We get that

$$f(y) - f(x_{-k}) \geq f(y) - \sum_{i \neq k} f(x_i) = f(x_k) - f(x_0).$$

If  $x_0 > x_{-k}$  then  $y + x_0 \geq \sum_{i=1}^n x_i$ . Otherwise, assume wlog that  $x_{-k} > x_k$ . By similar arguments to those used for  $n = 2$  we conclude. ■

## REFERENCES

- [1] [https://www.romande-energie.ch/images/Files/prix-electricite/2015\\_prix-electricite\\_RE.pdf](https://www.romande-energie.ch/images/Files/prix-electricite/2015_prix-electricite_RE.pdf), sampled on Sept 25, 2014.
- [2] Christopher Andrey and Alain Haurie. The economics of electricity dynamic pricing and demand response programmes. Technical report, ORDECSYS, 2013.
- [3] Stephen Poythress Boyd and Lieven Vandenbergh. *Convex optimization*. Cambridge university press, 2004.
- [4] Lijun Chen, Na Li, Steven H Low, and John C Doyle. Two market models for demand response in power networks. In *Smart Grid Communications (SmartGridComm)*, pages 397–402. IEEE, 2010.
- [5] Drew Fudenberg and Jean Tirole. *Game Theory*. MIT Press, Cambridge, MA, 1991.
- [6] Elena Marie Krieger. *Effects of variability and rate on battery charge storage and lifespan*. PhD thesis, Princeton University, 2013.
- [7] Jean-Yves Le Boudec and Dan-Cristian Tomozei. Demand Response Using Service Curves. In *IEEE ISGT-EUROPE*, 2011.
- [8] A.-H. Mohsenian-Rad, V.W.S. Wong, J. Jatskevich, R. Schober, and A. Leon-Garcia. Autonomous demand-side management based on game-theoretic energy consumption scheduling for the future smart grid. *Smart Grid, IEEE Transactions on*, 1(3):320–331, 2010.
- [9] Jean-Marc Oury and Bruno Heintz. Method and system of real time management and modulation of electricity consumption, January 29 2013. US Patent 8,364,322.
- [10] Thomas B Reddy. *Linden's Handbook of Batteries*, volume 4. McGraw-Hill, 2011.
- [11] Jörgen W Weibull. *Evolutionary game theory*. MIT press, 1997.